

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Toshihiko Munetsugu
Appl. No. : 10/733,981
Filed : December 11, 2003
Title : DATA PROCESSING DEVICE AND METHOD FOR SELECTING
MEDIA SEGMENTS ON THE BASIS OF A SCORE

Conf. No. : 1257
TC/A.U. : 2176
Examiner : Maikhanh Nguyen

Customer No. : 000,116
Docket No. : NGB-32161US2

SECOND SUPPLEMENT to APPELLANT'S BRIEF

Sir:

This paper is provided to supplement the brief provided in an ex parte appeal from the decision of the Examiner in the Final Rejection dated December 19, 2006 in the above-identified application, rejecting all claims in the application. This supplemental paper is provided in response to the Notice of Non-Compliant Appeal Brief mailed on April 28, 2008, and is meant to supplement the brief filed on December 19, 2007. Please extend the due date of May 28, 2008 by one month to June 30, 2008. The one-month extension of time fee of \$120 is paid herewith.

Remarks are provided on page 2; and

A supplemental **Summary of the Claimed Subject Matter** is attached hereto to supplement the brief already filed in this case.

REMARKS

Patent Appeals Specialist Doug Hutton sent a Notice of Non-Compliant Appeal Brief on April 28, 2008, which argues that the Summary of the Claimed Subject Matter is not compliant, arguing that “Appellant must amend the Summary of Claimed Subject Matter section to recite in exact language (i.e., quote) every limitation of the independent claims and identify each element of the limitations to those portions of the disclosure that describe it”, stating that the previously supplied summary fails to comply with 37 C.F.R. 41.37(c)(1)(v), because the ‘Summary of Claimed Subject Matter’ does not “map all independent claims on appeal, to the specification by page and line number of paragraph number and to the drawings, if any.”

Although a supplemental summary is provided with this response that complies with the specialist’s requested format, Applicant notes that the requested format is not based on the plain language of the rules, but is merely one manner in complying with the rules out of many different alternatives that would also comply with the rules. The rules do not require a *mapping* of the features of all of the claims to the specification is provided.

The requirements for the summary section of an appeal brief are laid out in 37 CFR 41.37(c)(1)(v) as follows:

Summary of claimed subject matter. A concise explanation of the subject matter defined in each of the independent claims involved in the appeal, which shall refer to the specification by page and line number, and to the drawing, if any, by reference characters. For each independent claim involved in the appeal and for each dependent claim argued separately under the provisions of paragraph (c)(1)(vii) of this section, every means plus function and step plus function as permitted by 35 U.S.C. 112, sixth paragraph, must be identified and the structure, material, or acts described in the specification as corresponding to each claimed function must be set forth with reference to the specification by page and line number, and to the drawing, if any, by reference characters.

(underlining and bold emphasis added). The rule only requires that a “concise explanation” of the “subject matter defined in each of the independent claims” be provided, and it is clear from reading this rule that the requirement that reference be made to the “specification by page and line number, and to the drawing if any” refers to the “concise explanation”. The definitions of “explanation” do not comport with quoting the “exact language” of the claims. Applicant notes that the rule specifically says that it is the summary which has the “concise explanation”, and

thus exactly quoting claim language can hardly be used to “explain” the claim language. One does not define a word by repeating the word, just as one does not properly “explain” claim language by merely repeating it verbatim.

Any reasonable reading of this rule requires only that (1) the subject matter of the independent claims are provided as part of the *concise explanation* and (2) this concise explanation should *refer* to the specification and drawings, where applicable. Nowhere does the rule require exact *quoting* of the claim language or any *mapping* of the claim language to the specification, and thus it is improper for the USPTO to impose such a more stringent rule upon Applicants when other alternatives would literally satisfy the rule. The previously provided summary section fully complied with this portion of the rule (there are references to both the specification and the drawings, where applicable).

Nevertheless, the newly supplied summary section also meets the literal requirements for the rule, along with the specialists *suggestions* (which are *not* part of the rule), and thus should also be acceptable.

If there are any additional fees resulting from this communication, please charge same to our Deposit Account No. 16-0820, our Order No. 32161US2.

Respectfully submitted,
PEARNE & GORDON, LLP

June 30, 2008

By: / Robert F. Bodi /
Robert F. Bodi, Reg. No. 48,540

1801 East Ninth Street
Suite 1200
Cleveland, Ohio 44114-3108
(216) 579-1700

SUMMARY OF THE CLAIMED SUBJECT MATTER

The invention provided by the subject matter found in independent claims 37, 47, 57 and 60 relates to a data processing apparatus for processing media content comprised of a plurality of scenes, as described in the clean copy of the replacement specification. In essence, the invention provides a new way of describing media content, through the use of context (content) description data that provides a contextual description of the media content (such as a video, for example). This description can then be used to select various scenes of the media content based on a user input, the scenes chosen based on importance related to a contextual topic as chosen by the user (see, e.g., the Summary of the Invention section of the specification).

Figure 1 (of the replacement drawings) shows a simplified view of a method performed by the apparatus, according to the invention. A simple description of this process is that, via a selection step 101, selected segments are determined according to context description data that is input into the apparatus (see the second paragraph of page 30, lines 9-17 of the clean version of the replacement specification for a description of the method shown in Figure 1). The selected segments are then used by the apparatus, as shown by the example in Figure 5, to select desired scenes of the media content, (e.g., audio/video) that is input into the apparatus (into a demultiplex means 601). The media content is split into audio data (input into an audio skimming means 603) and video data (input to a video skimming means 602). The selection segments are then utilized by the apparatus (e.g., the skimming means 602, 603) to output desired audio and video data based on the degree of importance of the scenes (see the last paragraph of page 34 of the specification, line 23 to the end of the paragraph on page 35 line 15).

The specification makes clear that context description data is *different* than the media content. On page 17 of the specification, last paragraph (page 17 line 23 to page 18, line 8), media content and context description data are differentiated. Media content is described as corresponding to *video data and/or audio data*, whereas context description data is described as the *configuration* of respective video data sets and/or audio data sets, and that video/audio selection means selects a scene by reference to the context description data, and that an extraction means then extracts the selected video or audio scene. On pages 30-33, context description data is described in more detail, as discussed below.

An example of the context description data is shown in Figure 2 in a hierarchical format, and is described on page 30, last paragraph, to page 31 (page 30, line 18 to page 31, line 24). Figures 19, 29, 36, 65, and 68 also show additional embodiments of this data, all of which are described in the specification. In this discussion, it is made clear that the context description data is data for *describing* the media content, such as the *context* of various scenes of the media content, but it is clear that this data is *not* the media content itself. For example, see the last paragraph of page 30 (lines 22-25), where it is stated that the context description data can be used to *describe* a movie story hierarchically, according to the movie, chapters, sections, and paragraphs. Similarly, a video of a baseball game could be *described* by using context description data by breaking down the game into innings, half-innings, at-bats, and individual pitches, for example, to describe the game (see end of page 31, line 25 to top of page 32, line 11). Thus, it is clear from the specification that context description data is differentiated from the media content that it describes.

On page 32 (starting at line 12), the example context description data of Figure 2 is described as potentially being expressed in XML language. An example of context description

data is provided on Program02.xml of the sample programs filed on a CD ROM in this case, and described on page 31. Such an XML file could, for example, be executed in a browser application on a computer (as described on page 31, for example). This example of context description data further clarifies that this data is not media content, but data that is provided to describe some media content. The XML ASCII file for Program02.xml is provided in the Evidentiary Appendix, attached to this brief. This file can be loaded into a standard browser, such as Microsoft's Internet Explorer, where it can be executed (however, an editorial error in this example program causes an error in the browser; such an error is well within the skill of the art to correct, however).

Figure 20 shows an input means (e.g., interface means 2401) and a selection means (item 1801) operable to input context description data (see Figure 19, discussed on page 52, line 8 to page 53, line 4) for performing the selection step of Figure 17 (see page 50, line 14 to page 51, line 12), which provides support for the "input means" and "selection means" of claim 57. The context description data can include a plurality of segment elements, such as shown in Figure 19, each for describing one of said plurality of scenes of media content (see page 52, line 8 to page 53, line 4), for input into the interface means 2401. The context description data can include a "context attribute", such as the keyword discussed on page 17 in the third paragraph, lines 17-22). The keyword describes a particular "viewpoint" (e.g., context) that the user desires (id.). See the data structure of Figure 29 showing keywords associated with a priority (a "score" provided in an importance attribute, discussed below). This is discussed in more detail on pages 61-63 and pages 32-34; see also page 49, line 15 to page 50, line 6).

The context description data also includes a plurality of importance attributes each associated with a corresponding one of the plurality of segment elements, for quantifying the impor-

tance (or “score”) of the particular desired context represented by an associated context attribute. This is shown by example in the priority elements of Figure 29, which are assigned a “priority” relating to the degree of importance of the scene, with these importance attributes having a value representing a degree of contextual importance of the corresponding scenes (see, e.g., page 63, lines 7-23). Data of the segments related to the scenes (e.g., scene start and stop times) are then output based on the one or more importance attributes (see e.g., page 63, line 14 to page 65, line 1). In this manner, context description data can be used to determine and select scenes having a high degree of contextual importance related to a context attribute (e.g., keyword) of a user request (id.; also see, e.g., Fig. 43 and pages 89-91 of the specification for more detail about user requests).

The Evidentiary appendix includes an example of the context description data for the above described process (in Program02.xml), along with an example output (Program07.out) showing the start/stop times for the corresponding scenes.

In particular, Claim 37 recites a **data processing apparatus** for processing media content comprised of a plurality of scenes, with the apparatus comprising the following elements, mapped to the specification and drawings:

an input unit operable to input content description data (see, e.g., selection means discussed on page 16, lines 8-17; See also Fig. 20, interface means 2401 connected to database 1805, and selection means 1801, discussed on page 53, line 14 to line 3 of page 55) including a plurality of segments each for describing one of said plurality of scenes of media content (see, e.g., Figs 2 discussed on page 30, lines 18-25), said content description data further including:

a context attribute having a value for describing a context of said media content (e.g., a “keyword” page 17, lines 17-22; see also page 62, line 25 and page 65, lines 3-11);

a plurality of importance attributes each associated with one of said segments and having a value representing a degree of contextual importance of said corresponding one of said segments (e.g., the “score”, see page 17, lines 11-22, and “priority: shown in Figures 2 and 3, discussed on last paragraph of page 31 lines 1-6 and 12-18);

and

an output unit operable to output at least one of said segments based on at least one of said importance attributes (see, e.g., Fig. 20, demultiplex means 2402, video skimming means 2403, and audio skimming means 2404, and first paragraph of page 54, line 4, continuing to page 56, line 8, discussing Fig. 21).

Furthermore, Claim 47 recites a **data processing method** for processing media content comprised of a plurality of scenes, with the method comprising the following steps, mapped to the specification and drawings:

inputting content description data (see e.g., selection step 101 of Figure 1, discussed on page 30, lines 9-25) including a plurality of segments each for describing one of said plurality of scenes of media content (see, e.g., Figs 2 discussed on page 30, lines 18-25), said content description data including:

a context attribute having a value for describing a context of said media content (e.g., “keyword” page 17, lines 17-22; see also page 62, line 25 and page 65, lines 3-11),
and

a plurality of importance attributes each associated with one of said segments and having a value representing a degree of contextual importance of said corresponding one

of said segments(e.g., the “score”, see page 17, lines 11-22, and “priority: shown in Figures 2 and 3, discussed on last paragraph of page 31 lines 1-6 and 12-18);

and

outputting at least one of said segments based on at least one of said importance attributes (see, e.g., Fig. 20, demultiplex means 2402, video skimming means 2403, and audio skimming means 2404, and first paragraph of page 54, line 4, continuing to page 56, line 8, discussing Fig. 21).

In addition, Claim 57 recites a **data processing apparatus** for processing media content comprised of a plurality of scenes, with the apparatus comprising the following elements, mapped to the specification and drawings:

input means for inputting (see, e.g., interface means 2401 connected to database 1805, discussed on the first full paragraph of page 54, line 4 to line 3 of page 55) content description data describing said plurality of scenes (see, e.g., Figs 2 discussed on page 30, lines 18-25), said content description data being arranged in a hierarchy and including:

a plurality of section elements (see, e.g., fig. 29 discussed at page 63, line 6 to page 65, line 1);

a plurality of segment elements each being a child of one of said section elements and also being associated with a corresponding one of said plurality of scenes (id.);

a plurality of context attributes (see “keyword” page 17, lines 17-22; see also of Fig. 29, page 62, line 25 and page 65, lines 3-11) each being associated with one or more of said segment elements and/or section elements (id.), each of said context attributes having a value for describing a context of said media content (id.);

a plurality of importance attributes each associated with one of said context attributes and also associated with one of said segment elements that are associated with said one of said context attributes, and having a value representing a degree of impor-

tance of the scene associated with said one of said segment elements in relation to the context of said context attribute (e.g., the “score”, see page 17, lines 11-22, and “priority: shown in Figures 2 and 3, discussed on last paragraph of page 31 lines 1-6 and 12-18; and see Fig. 29 discussed above), and

a plurality of time attributes (page 64, line 11 to page 65, line 1) each associated with a corresponding one of said segments for determining a start time and one of an end time (id.) and a duration of the scene associated with said corresponding segment (id.); and

selection means (item 1801 of Fig. 20, discussed on page 54, lines 4-15) for selecting one or more of said segments (id.) based on an analysis of said importance attributes (page 63, line 13 to page 65, line 1).

Finally, Claim 60 recites a **data processing method** for processing media content comprised of a plurality of scenes, with the method comprising the following steps, mapped to the specification and drawings:

Inputting (see, e.g., interface means 2401 connected to database 1805, discussed on the first full paragraph of page 54, line 4 to line 3 of page 55) hierarchically arranged context description data that describes a plurality of scenes of the media contents of one or more media files (see, e.g., Figs 2 discussed on page 30, lines 18-25), said context description data including:

a plurality of segment elements each for describing one of said plurality of scenes (see, e.g., fig. 29 discussed at page 63, line 6 to page 65, line 1),

a plurality of section elements each having either one or more of said plurality of section elements as children, or having one or more of said plurality of segment elements as children (id.),

a plurality of context attributes each having a value for describing a corresponding context of said media content and each being an attribute associated with one or more of said segment elements and including at least one keyword for describing the contents of the scenes described by the associated one or more of said segment elements (see, e.g., “keyword” page 17, lines 17-22; see also of Fig. 29, page 62, line 25 and page 65, lines 3-11), and

a plurality of importance attributes each associated with a corresponding one of said segment elements and having a value representing a degree of importance of the scene corresponding to said corresponding segment element in relation to one context attribute that is also associated with corresponding segment element (e.g., the “score”, see page 17, lines 11-22, and “priority: shown in Figures 2 and 3, discussed on last paragraph of page 31 lines 1-6 and 12-18; and see Fig. 29 discussed above);

selecting one or more of said segment elements based on an analysis of one or more of said context attributes and the associated importance attributes (e.g., item 1801 of Fig. 20, discussed on page 54, lines 4-15 and page 63, line 13 to page 65, line 1);

inputting said media content (see, e.g., Figs. 1, 5, and 11, among others; and the specification at page 30, lines 9-18; page 34 line 23 to page 35 line 15; and page 41, lines 2-11, respectively)

outputting one or more of said plurality of scenes based on the selected segment elements (id.).

In this manner, the invention as claimed provides a new and useful method and apparatus for selecting, playing back, delivering a synopsis, highlighting a scene, and/or selecting a scene desired by the audience at the time of playback of the media content (see first paragraph of the Summary section of the specification).